

# Northumbria Research Link

Citation: Shang, Yilun (2019) Attack robustness and stability of generalized k-cores. New Journal of Physics, 21. ISSN 1367-2630

Published by: IOP Publishing

URL: <https://iopscience.iop.org/article/10.1088/1367-2630/ab3d7c>  
<<https://iopscience.iop.org/article/10.1088/1367-2630/ab3d7c>>

This version was downloaded from Northumbria Research Link:  
<http://nrl.northumbria.ac.uk/id/eprint/40639/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



**Northumbria  
University**  
NEWCASTLE



**UniversityLibrary**



## PAPER

Attack robustness and stability of generalized  $k$ -coresYilun Shang 

Department of Computer and Information Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, United Kingdom

E-mail: [yilun.shang@northumbria.ac.uk](mailto:yilun.shang@northumbria.ac.uk)**Keywords:** complex network, robustness, localized attack, Erdős–Rényi network, exponential network, stability, targeted attacks (TA)

## OPEN ACCESS

RECEIVED  
15 June 2019REVISED  
9 August 2019ACCEPTED FOR PUBLICATION  
21 August 2019PUBLISHED  
10 September 2019

Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/4.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



## Abstract

Earlier studies on network robustness have mainly focused on the integrity of functional components such as the giant connected component in a network. Generalized  $k$ -core ( $Gk$ -core) has been recently investigated as a core structure obtained via a  $k$ -leaf removal procedure extending the well-known leaf removal algorithm. Here, we study analytically and numerically the network robustness in terms of the numbers of nodes and edges in  $Gk$ -core against random attacks (RA), localized attacks (LA) and targeted attacks (TA), respectively. In addition, we introduce the concept of  $Gk$ -core stability to quantify the extent to which the  $Gk$ -core of a network contains the same nodes under independent multiple RA, LA and TA, respectively. The relationship between  $Gk$ -core robustness and stability has been studied under our developed percolation framework, which is of significance in better understanding and design of resilient networks.

## 1. Introduction

In network science, robustness refers to the ability of surviving random failures or intentional attacks. Much work has been carried out to explore the robustness of networked systems by revealing the size of their functional components through percolation theory [1, 2]. In this context, the most studied functional components in networks include giant connected component [3, 4],  $k$ -core structure [5–7] and core [8–10]. A recent study [11] considered a  $k$ -leaf removal procedure for  $k \geq 2$  which leads to the Generalized  $k$ -core (or  $Gk$ -core) by progressively removing  $k$ -leaves, i.e. nodes with degree less than  $k$ , together with their nearest neighbors and all incident edges. It is clear that the resulting subgraph is equivalent to the ordinary core in the case of  $k = 2$  [8]. The  $Gk$ -core naturally characterizes for example the robustness of networks suffering from virus infection deactivating weak nodes (i.e.  $k$ -leaves) and their nearest neighbors.

However, in most real situations, such as power grid blackouts, market crashes, and brain seizures, understanding merely the network robustness in terms of the size of functional component is not very useful. The potential damage and recovery of the network crucially rely on the location of the functional components. [12] showed a pronounced variation of giant component sizes corresponding to two correlated random realizations of percolation, suggesting the changing role of individual nodes in response to different percolation. Stability is recently introduced in [13] as a novel measure to quantify the extent to which the giant connected component of a network consists of the same nodes under multiple independent edge percolation. It is found interestingly that robustness and stability are consistent in single-layer networks but do not always imply each other in networks with interdependency links.

Here we extend the stability of giant components under independent edge percolation to stability of  $Gk$ -core under a range of attacks including random attacks (RA) [14–17], localized attacks (LA) [18–20], and targeted attacks (TA) [21–23]. Moreover, we investigate how each type of attack influences the network robustness in terms of the numbers of nodes and edges in the  $Gk$ -cores. It is found that the effect of a LA is exactly the same as that of a RA on Erdős–Rényi (ER) networks in terms of both robustness and stability of  $Gk$ -core. Interestingly, the analogous equivalence recurs in exponential networks but between LA and TA (see table 1). Under our percolation framework, we observe discontinuous percolation transition for  $Gk$ -core with  $k \geq 3$  and continuous percolation transition for  $G2$ -core in all attack scenarios. The relationship between robustness

**Table 1.** The robustness and stability of  $Gk$ -core under RA, LA and TA for given  $k$ . For example, in ER networks, ‘RA = LA > TA’ means that the network is more robust (stable) under RA and LA than under TA; the robustness (stability) under RA and LA are equivalent.

	Erdős–Rényi network	Random regular network	Exponential network
Robustness	RA = LA > TA	LA > RA = TA	RA > LA = TA
Stability	RA = LA > TA	LA > RA = TA	RA > LA = TA

and stability is explored in three stylized network models as well as a real-world social network. We find excellent agreement between theoretical calculations and numerical simulations.

The rest of the work is organized as follows. The analytical frameworks for attack robustness and stability based on generating function formalism are established in section 2. Numerical studies for synthetic networks and an example of a large-scale real-life network are given in sections 3 and 4, respectively. The conclusion is drawn in section 5.

## 2. Theoretical results

In this section, we consider a random network model with any degree distribution of node degree. Specifically, let  $P(q)$  be the probability that a randomly chosen node has degree  $q$ . Let  $n$  and  $l$  be the numbers of nodes and edges in the network, respectively. Following [3, 24], the generating function for the degree distribution is defined by  $G_0(x) = \sum_q P(q)x^q$  and  $G_1(x) = G'_0(x)G'_0(1)^{-1}$  is the generating function for the excess degree distribution. Clearly,  $2l/n = G'_0(1)$  is the average degree. We are interested in the numbers of nodes and edges as well as stability of  $Gk$ -core when a fraction  $1 - p$  of nodes are removed according to RA, LA, and TA.

Under RA, we define the stability of  $Gk$ -core as the fraction of nodes in all  $Gk$ -cores under  $\ell$  independent realizations of random attacks on the network. Namely,

$$S_k^{\text{RA}}(\ell, p) := \frac{1}{n} |\bigcap_{t=1}^{\ell} Gk\text{-core}_t|, \quad (1)$$

where  $Gk\text{-core}_t$  means the nodes in  $Gk$ -core in the  $t$ th realization of RA (in which a fraction  $1 - p$  of nodes are randomly removed), and  $|\cdot|$  means the size of a set. We will omit the superscript or the parameters  $\ell, p$  in (1) (and other similar notations later) when no confusion will be caused. For LA and TA, we have similar definitions for the corresponding stability. Equation (1) extends the stability concept of giant component [13] and characterizes the extent to which the  $Gk$ -core is stable regardless of the specific damage caused during an attack.

For a given  $p$ , if the expected size of  $Gk$ -core under consideration is denoted by  $m$ , it would be useful to compare  $S_k(\ell, p)$  with the ‘stability’ of a random subset:

$$S(\ell) := \frac{1}{n} |\bigcap_{t=1}^{\ell} R_t|, \quad (2)$$

where  $R_t$  is the  $t$ th realization of a random set of size  $m$  sampled randomly from the network with replacement. Note that  $|\bigcap_{t=1}^{\ell} R_t|$  follows the binomial distribution  $\text{Bin}(n, (m/n)^\ell)$  with expectation  $n(m/n)^\ell$ . Hence, the expected  $S(\ell)$  decays exponentially in the form  $(m/n)^\ell$  as a function of  $\ell$ .

### 2.1. Generalized $k$ -core under random attacks

A random attack removing a fraction  $1 - p$  of nodes from the network can be considered as a two-staged process by first removing the nodes but keeping the edges connecting the remaining nodes and the removed nodes, and then removing those edges. In other words, in the first stage only nodes are deleted and in the second stage only edges are deleted. Since the network is randomly connected, the probability of a random edge leaving a removed node is equal to the ratio of the number of edges leaving the removed nodes in the first stage to the total number of edges leaving all nodes in the original network. Hence, the probability for an edge to leave a removed node can be calculated as

$$\frac{2l - np \cdot \frac{2l}{n}}{2l} = 1 - p. \quad (3)$$

The generating function of the degree distribution, denoted by  $\hat{G}_0(x)$ , of the resulting network after the random attack becomes

$$\hat{G}_0(x) = G_0(1 - p + px) := \sum_q \hat{P}(q)x^q \quad (4)$$

following [16] because deleting the edges leaving the removed nodes is equivalent to deleting a  $1 - p$  fraction of edges randomly in the second stage.

Given  $k \geq 2$ , the  $Gk$ -core is obtained by a iterative pruning procedure. At each time step, a randomly chosen  $k$ -leaf is removed together with its nearest neighbors and all their incident edges. The procedure continues until no  $k$ -leaves exist in the remaining network. The resulting subgraph is called the  $Gk$ -core. Following the approach of [11], the nodes can be split into three categories: if a node can become a leaf, it is called  $\alpha$ -removable; if a node can become a neighbor of a leaf, it is called  $\beta$ -removable; if a node belongs to  $Gk$ -core, it is called non-removable. Assume that following a randomly selected edge, the node we arrive at is  $\alpha$ -removable,  $\beta$ -removable, and non-removable with probability  $\alpha$ ,  $\beta$ , and  $1 - \alpha - \beta$ , respectively. Invoking (4), these probabilities can be calculated as [11]

$$\alpha = \frac{1}{\hat{G}'_0(1)} \sum_{s=0}^{k-2} \frac{(1 - \alpha - \beta)^s}{s!} \hat{G}_0^{(s+1)}(\beta) = \frac{1}{G'_0(1)} \sum_{s=0}^{k-2} \frac{(1 - \alpha - \beta)^s p^s}{s!} G_0^{(s+1)}(1 - p + \beta p), \quad (5)$$

$$\beta = 1 - \frac{\hat{G}'_0(1 - \alpha)}{\hat{G}'_0(1)} = 1 - \frac{G'_0(1 - \alpha p)}{G'_0(1)}, \quad (6)$$

and the relative size of  $Gk$ -core, denoted by  $n_k^{\text{RA}}(p)$ , is [11]

$$\begin{aligned} n_k^{\text{RA}}(p) &= p \hat{G}_0(1 - \alpha) - p \sum_{s=0}^{k-1} \frac{(1 - \alpha - \beta)^s}{s!} \hat{G}_0^{(s)}(\beta) \\ &= p G_0(1 - \alpha p) - \sum_{s=0}^{k-1} \frac{(1 - \alpha - \beta)^s p^{s+1}}{s!} G_0^{(s)}(1 - p + \beta p), \end{aligned} \quad (7)$$

where the notation  $G_0^{(s)}$  stands for the  $s$ th derivative of  $G_0$ . The normalized number of edges in  $Gk$ -core, signified by  $l_k^{\text{RA}}(p)$ , can be computed by (3) as

$$l_k^{\text{RA}}(p) = (1 - \alpha - \beta)^2 \frac{p^2 l}{n} = (1 - \alpha - \beta)^2 \frac{p^2}{2} G'_0(1), \quad (8)$$

where  $(1 - \alpha - \beta)^2$  is the probability that both end nodes of a random edge belong to the  $Gk$ -core, and  $p^2 l$  corresponds to the numbers of edges in the network after random attack. Note that  $l_k^{\text{RA}}(p)$  can also be derived by using generating function (4).

Next, we study the stability  $S_k^{\text{RA}}(\ell, p)$  for  $\ell \geq 1$ . Note that a node is in the  $Gk$ -core if it has at least  $k$  neighbors which are also in the  $Gk$ -core. By using (1), (4), and the repeated differentiation of the generating function [3], we have the expected stability

$$\begin{aligned} S_k^{\text{RA}}(\ell, p) &= p \sum_{q \geq k} \hat{P}(q) \left[ \sum_{s=k}^q \binom{q}{s} (1 - \alpha - \beta)^s \beta^{q-s} \right]^\ell \\ &= \sum_{q \geq k} \frac{p^{q+1} G_0^{(q)}(1 - p)}{q!} \left[ (1 - \alpha)^q - \sum_{s=0}^{k-1} \frac{q! (1 - \alpha - \beta)^s \beta^{q-s}}{(q-s)! s!} \right]^\ell, \end{aligned} \quad (9)$$

where the first factor  $p$  is the occupation probability of RA,  $\hat{P}(q)$  is the degree distribution after RA,  $\binom{q}{s}$  is the binomial coefficient, and the term in the brackets is the probability that a randomly chosen node belongs to  $Gk$ -core given that it has degree  $q$ . The last equality in (9) follows directly from the differentiation property as well as the binomial theorem. Note that a one line calculation validates  $S_k^{\text{RA}}(1, p) = n_k^{\text{RA}}(p)$  in the light of (7) and (9).

## 2.2. Generalized $k$ -core under localized attacks

A localized attack is performed by randomly deleting a seed node in the network, then its nearest neighbors, and then its second nearest neighbors and so on until a fraction  $1 - p$  of nodes in the whole network are removed. The generating function of the degree distribution, denote by  $\hat{G}_0(x)$ , of the resulting network after the localized attack is given by [18]

$$\hat{G}_0(x) = \frac{1}{G_0(f)} G_0 \left( f + \frac{G'_0(f)}{G'_0(1)} (x - 1) \right), \quad (10)$$

where  $f = G_0^{-1}(p)$ . Likewise, we can write  $\hat{G}_0(x) = \sum_q \hat{P}(q) x^q$  with  $\hat{P}(q) = (q!)^{-1} \hat{G}_0^{(q)}(0)$ .

Using (10) and following the approach of [11], the probabilities  $\alpha$  and  $\beta$  after  $Gk$ -core percolation are established by

$$\begin{aligned}\alpha &= \frac{1}{\hat{G}'_0(1)} \sum_{s=0}^{k-2} \frac{(1-\alpha-\beta)^s}{s!} \hat{G}_0^{(s+1)}(\beta) \\ &= \frac{G'_0(1)}{G'_0(f)^2} \sum_{s=0}^{k-2} \frac{(1-\alpha-\beta)^s G'_0(f)^{s+1}}{s! G'_0(1)^{s+1}} G_0^{(s+1)} \left( f + \frac{G'_0(f)}{G'_0(1)} (\beta-1) \right),\end{aligned}\quad (11)$$

$$\beta = 1 - \frac{\hat{G}'_0(1-\alpha)}{\hat{G}'_0(1)} = 1 - \frac{G'_0 \left( f - \frac{G'_0(f)}{G'_0(1)} \alpha \right)}{G'_0(f)}, \quad (12)$$

and the relative size of  $Gk$ -core, denoted by  $n_k^{\text{LA}}(p)$ , is

$$\begin{aligned}n_k^{\text{LA}}(p) &= p \hat{G}_0(1-\alpha) - p \sum_{s=0}^{k-1} \frac{(1-\alpha-\beta)^s}{s!} \hat{G}_0^{(s)}(\beta) \\ &= G_0 \left( f - \frac{G'_0(f)}{G'_0(1)} \alpha \right) - \sum_{s=0}^{k-1} \frac{(1-\alpha-\beta)^s G'_0(f)^s}{s! G'_0(1)^s} \cdot G_0^{(s)} \left( f + \frac{G'_0(f)}{G'_0(1)} (\beta-1) \right),\end{aligned}\quad (13)$$

where again  $f = G_0^{-1}(p)$ . The normalized number of edges in  $Gk$ -core, denoted by  $l_k^{\text{LA}}(p)$ , can be computed by (10) similarly as in (8):

$$l_k^{\text{LA}}(p) = p(1-\alpha-\beta)^2 \frac{\hat{G}'_0(1)}{2} = (1-\alpha-\beta)^2 \frac{G'_0(f)^2}{2G'_0(1)}, \quad (14)$$

where  $\hat{G}'_0(1)$  is the average degree of the network after localized attack.

To find the stability of  $Gk$ -core under LA, we argue similarly as in the RA scenario. It follows from (1) and (10) that

$$\begin{aligned}S_k^{\text{LA}}(\ell, p) &= p \sum_{q \geq k} \hat{P}(q) \left[ \sum_{s=k}^q \binom{q}{s} (1-\alpha-\beta)^s \beta^{q-s} \right]^\ell \\ &= \sum_{q \geq k} \frac{G'_0(f)^q G_0^{(q)} \left( f - \frac{G'_0(f)}{G'_0(1)} \alpha \right)}{q! G'_0(1)^q} \cdot \left[ (1-\alpha)^q - \sum_{s=0}^{k-1} \frac{q! (1-\alpha-\beta)^s \beta^{q-s}}{(q-s)! s!} \right]^\ell,\end{aligned}\quad (15)$$

where the term in the brackets is the probability that a randomly chosen node belongs to  $Gk$ -core given that it has degree  $q$ . It can be directly checked that  $S_k^{\text{LA}}(1, p) = n_k^{\text{LA}}(p)$  by (13) and (15).

### 2.3. Generalized $k$ -core under targeted attacks

A targeted attack removing a fraction  $1-p$  of nodes from the network can be conveniently implemented by assigning weight or probability to each node. For a node  $i$  with degree  $q_i$ , we set the canonical removal probability as

$$W_\gamma = \frac{q_i^\gamma}{\sum_{i=1}^n q_i^\gamma}, \quad (16)$$

for some real parameter  $\gamma \in (-\infty, +\infty)$ . When  $\gamma > 0$ , a node with higher degree is more likely to be deleted; when  $\gamma < 0$ , a node with lower degree instead is more likely to be deleted. The case  $\gamma \rightarrow \infty$  corresponds to the intentional deletion according to the fully sorted degree sequence. In particular,  $\gamma = 0$  is equivalent to the RA where each node is deleted with equal probability. Following [21], by introducing an auxiliary generating function  $\tilde{G}_0(x) = p^{-1} \sum_q P(q) t^{q^\gamma} x^q$  satisfying  $t = G_\gamma^{-1}(p)$  and  $G_\gamma(x) = \sum_q P(q) x^{q^\gamma}$ , we obtain the generating function for the degree distribution of the resulting network after target attack as

$$\hat{G}_0(x) = \tilde{G}_0(1 - \tilde{p} + \tilde{p}x) := \sum_q \hat{P}(q) x^q, \quad (17)$$

where  $\tilde{p} = (\sum_q P(q) q t^{q^\gamma}) / (\sum_q P(q) q)^{-1}$ .

Likewise, using (17) and following the approach of [11], the probabilities  $\alpha$  and  $\beta$  after  $Gk$ -core percolation are established by

$$\alpha = \frac{1}{\hat{G}'_0(1)} \sum_{s=0}^{k-2} \frac{(1-\alpha-\beta)^s}{s!} \hat{G}_0^{(s+1)}(\beta) = \frac{1}{\tilde{G}'_0(1)} \sum_{s=0}^{k-2} \frac{(1-\alpha-\beta)^s \tilde{p}^s}{s!} \tilde{G}_0^{(s+1)}(1 - \tilde{p} + \tilde{p}\beta), \quad (18)$$

$$\beta = 1 - \frac{\hat{G}'_0(1-\alpha)}{\hat{G}'_0(1)} = 1 - \frac{\tilde{G}'_0(1-\alpha\tilde{p})}{\tilde{G}'_0(1)}, \quad (19)$$

and the relative size of  $Gk$ -core, denoted by  $n_k^{\text{TA}}(p)$ , is

$$\begin{aligned} n_k^{\text{TA}}(p) &= p\hat{G}_0(1 - \alpha) - p \sum_{s=0}^{k-1} \frac{(1 - \alpha - \beta)^s}{s!} \hat{G}_0^{(s)}(\beta) \\ &= p\tilde{G}_0(1 - \alpha\tilde{p}) - p \sum_{s=0}^{k-1} \frac{(1 - \alpha - \beta)^s \tilde{p}^s}{s!} \tilde{G}_0^{(s)}(1 - \tilde{p} + \beta\tilde{p}). \end{aligned} \quad (20)$$

The normalized number of edges in  $Gk$ -core, denoted by  $l_k^{\text{TA}}(p)$ , can be computed similarly:

$$l_k^{\text{TA}}(p) = p(1 - \alpha - \beta)^2 \frac{\hat{G}'_0(1)}{2} = p(1 - \alpha - \beta)^2 \frac{\tilde{p}}{2} \tilde{G}'_0(1), \quad (21)$$

where  $\hat{G}'_0(1)$  is the average degree of the network after targeted attack.

Finally, we consider the stability  $S_k^{\text{TA}}(\ell, p)$  for  $\ell \geq 1$ . Note that a node is in the  $Gk$ -core if it has at least  $k$  neighbors which are also in the  $Gk$ -core. By using (1), (17), and the differentiation property of the generating function [3], we obtain the expected stability

$$\begin{aligned} S_k^{\text{TA}}(\ell, p) &= p \sum_{q \geq k} \hat{P}(q) \left[ \sum_{s=k}^q \binom{q}{s} (1 - \alpha - \beta)^s \beta^{q-s} \right]^\ell \\ &= p \sum_{q \geq k} \frac{\tilde{p}^q \tilde{G}_0^{(q)}(1 - \tilde{p})}{q!} \left[ (1 - \alpha)^q - \sum_{s=0}^{k-1} \frac{q! (1 - \alpha - \beta)^s \beta^{q-s}}{(q-s)! s!} \right]^\ell, \end{aligned} \quad (22)$$

where, as in RA and LA cases, the term in the brackets above is the probability that a randomly chosen node belongs to  $Gk$ -core given that it has degree  $q$ . It follows from (20) and (22) that  $S_k^{\text{TA}}(1, p) = n_k^{\text{TA}}(p)$  confirming the idea that stability at  $\ell = 1$  is equivalent to node fraction of  $Gk$ -core. Moreover, it is easy to check that the case of RA can be recovered as  $p = \tilde{p}$  and  $G_0(x) = \tilde{G}_0(x)$  when  $\gamma = 0$ .

### 3. Synthetic networks

We apply the above analytical framework for attack robustness and stability to three types of complex networks: homogeneous random networks following Poisson degree distributions and degenerate degree distributions, and quasi-heavy tailed networks following exponential degree distributions. Numerical simulations are based on networks with  $n = 10^7$  nodes. Here, we leave scale-free networks off since they only have a trivial  $Gk$ -core for all  $k \geq 2$  [11]. In table 1, we summarize the robustness and stability for  $Gk$ -core under RA, LA, and TA for these benchmark networks.

#### 3.1. ER networks

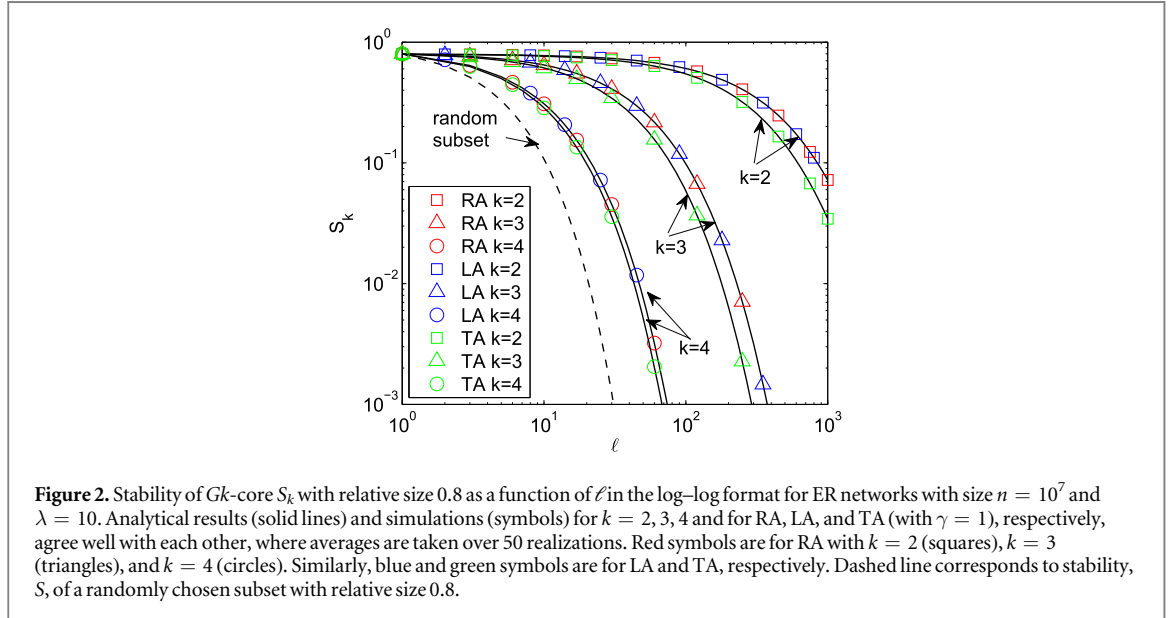
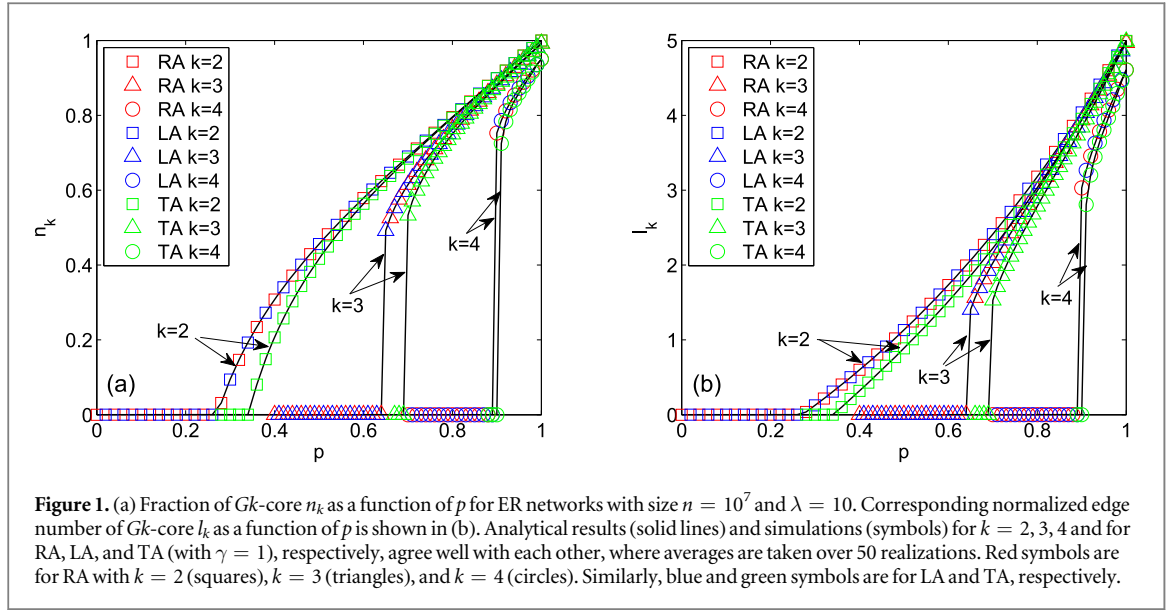
For an ER network with average degree  $\lambda$ , the degree distribution follows  $P(q) = e^{-\lambda} \lambda^q / q!$  for  $q \geq 0$ , and  $G_0(x) = e^{\lambda(x-1)}$ . In figure 1 we display the behavior of  $n_k(p)$  and  $l_k(p)$  as functions of occupation fraction  $p$  for ER networks with  $\lambda = 10$  under RA, LA, and TA with  $\gamma = 1$ . The agreement between simulations and results from generating function formalism is good. Some interesting observations are commented as follows.

The effect of an RA is exactly the same as that of an LA in terms of both  $n_k$  and  $l_k$ . In fact, we have  $n_k^{\text{RA}}(p) = n_k^{\text{LA}}(p)$  and  $l_k^{\text{RA}}(p) = l_k^{\text{LA}}(p)$  using some simple calculations with equations (7), (8), (13), and (14). This coincidence for ER networks has been observed for giant connected components in [18–20, 25] and for cores in [26]. This can be intuitively explained as a competition between degree heterogeneity, where high-degree nodes are more likely to sit in the attack hole of LA, and localization, where only surface nodes of the hole are connected to the remaining nodes. These two competitive forces of LA reach a balance in ER networks giving rise to the same damage measured by  $Gk$ -cores as an RA does.

With the variation of occupation probability  $p$ , continuous phase transition is observed for  $k = 2$  for all three types of attacks, while there is first order percolation transition behavior for  $k \geq 3$  in all attack scenarios (see figure 1). This is similar to  $k$ -core percolation [26].

When comparing damage caused to the  $Gk$ -core under TA with that under RA or LA, we find that TA is more harmful as high-degree nodes tend to be deleted in early stages dismantling the  $Gk$ -cores. From both figures 1(a) and (b), we note that this influence however gets smaller when  $k$  increases. This phenomenon can be explained as  $k$ -leaves for larger  $k$  are more likely to connect to some high-degree nodes. Therefore, removing these  $k$ -leaves will lead to the deletion of high-degree nodes. This effect turns out to be comparable to the TA with  $\gamma = 1$  considered here in ER networks for  $k = 4$ .

Next, we explore the relationship between  $Gk$ -core stability of ER networks under RA, LA, and TA. Several interesting observations can be derived from figure 2, where  $S_k(p, \ell)$  is shown for  $Gk$ -core with relative size 0.8 under all three types of attacks. Firstly,  $S_2 > S_3 > S_4 > S = 0.8^\ell$  (see equation (2)) for any given  $\ell$  under all



attack scenarios. This means the inner  $G_k$ -core with larger  $k$  tends to be less stable, which agrees with our intuition as the more nodes ( $k$ -leaves and their neighbors) are deleted the more fluctuation will be introduced. All  $S_k$  seems to decay exponentially but at slower rates than the random subset scenario. Secondly, we have  $S_k^{\text{RA}} = S_k^{\text{LA}} > S_k^{\text{TA}}$  for all  $k$  and  $\ell$ . The equivalence of  $S_k^{\text{RA}}$  and  $S_k^{\text{LA}}$  can be shown directly by using (9) and (15). Under TA, we are more likely to remove high-degree nodes, which are often ‘anchor nodes’ [13] in  $G_k$ -cores. Thus,  $S_k^{\text{TA}}$  drops lower than  $S_k^{\text{RA}}$  for any given  $k$ .

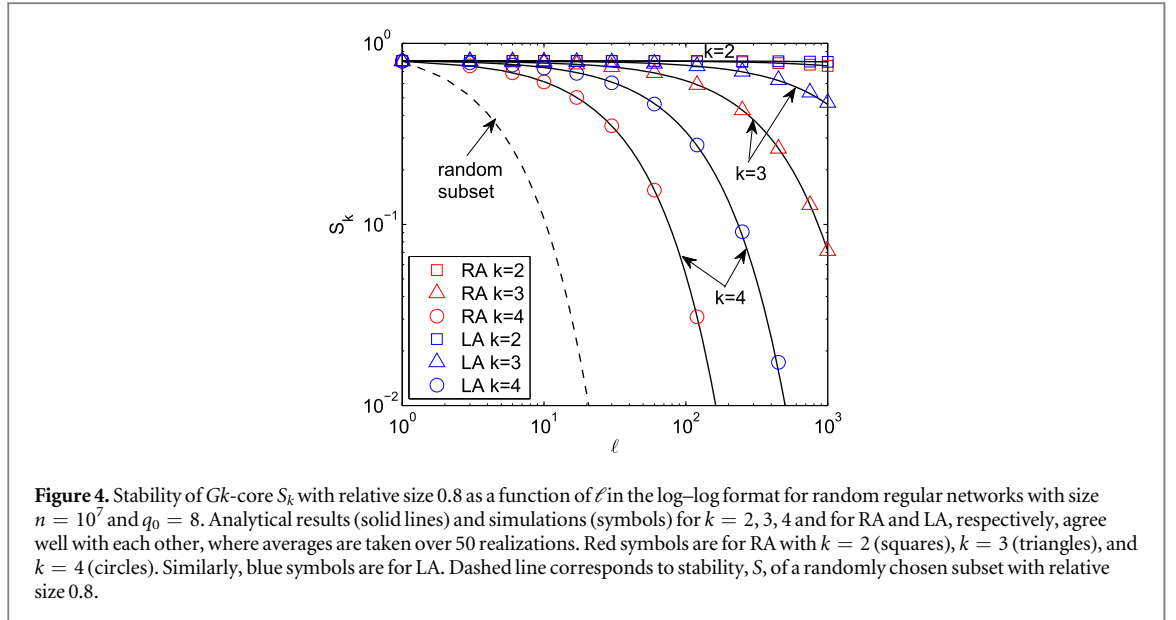
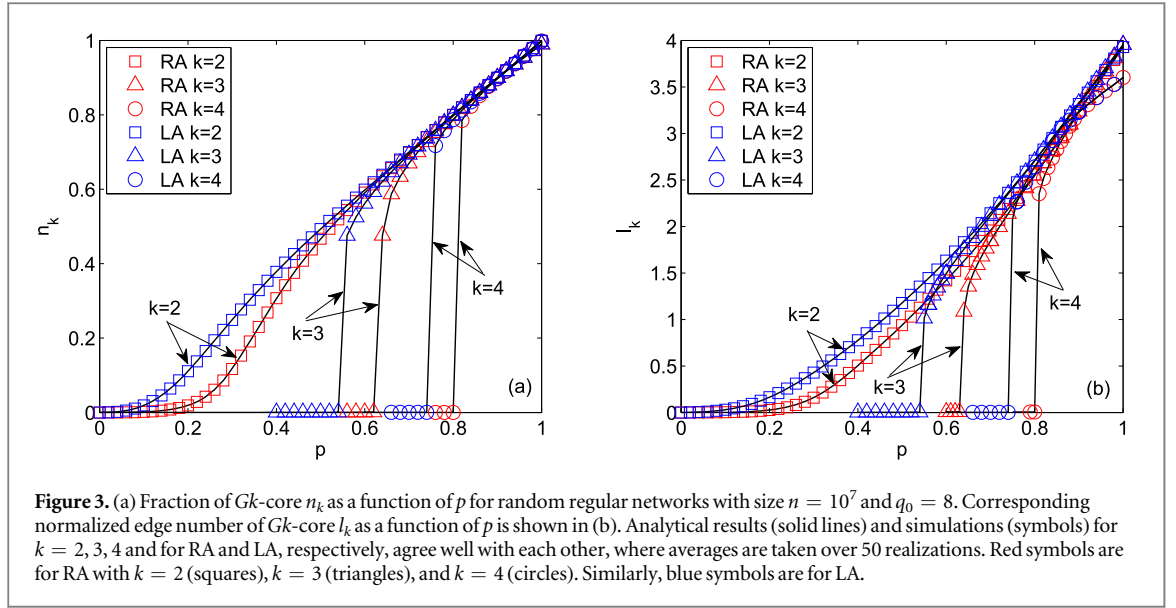
We can conclude from figures 1 and 2 that attack robustness and stability of  $G_k$ -cores for ER networks are generally consistent: lower  $k$  and milder attacks result in more robust and stable  $G_k$ -cores; see table 1. The stability of  $G_k$ -core has been shown in figure 2 (and below in figures 4 and 6) for the relative size of 0.8 as an example. It has qualitatively the same behavior for other relative sizes.

### 3.2. Random regular networks

For a random regular network degenerated on the atomic degree at  $q_0$ , the degree distribution follows  $P(q) = \delta_{q,q_0}$  for  $q \geq 0$ , and  $G_0(x) = x^{q_0}$ . Figures 3 and 4 correspond to the network of  $q_0 = 8$ . Noting that the two strategies RA and TA coincide for any value of  $\gamma$  since the nodes have the same degree in the initial network.

Similar as in ER networks, continuous phase transition is observed for  $k = 2$  for all attack scenarios, and there is first order percolation transition behavior for  $k \geq 3$  for all attack scenarios (see figure 3).





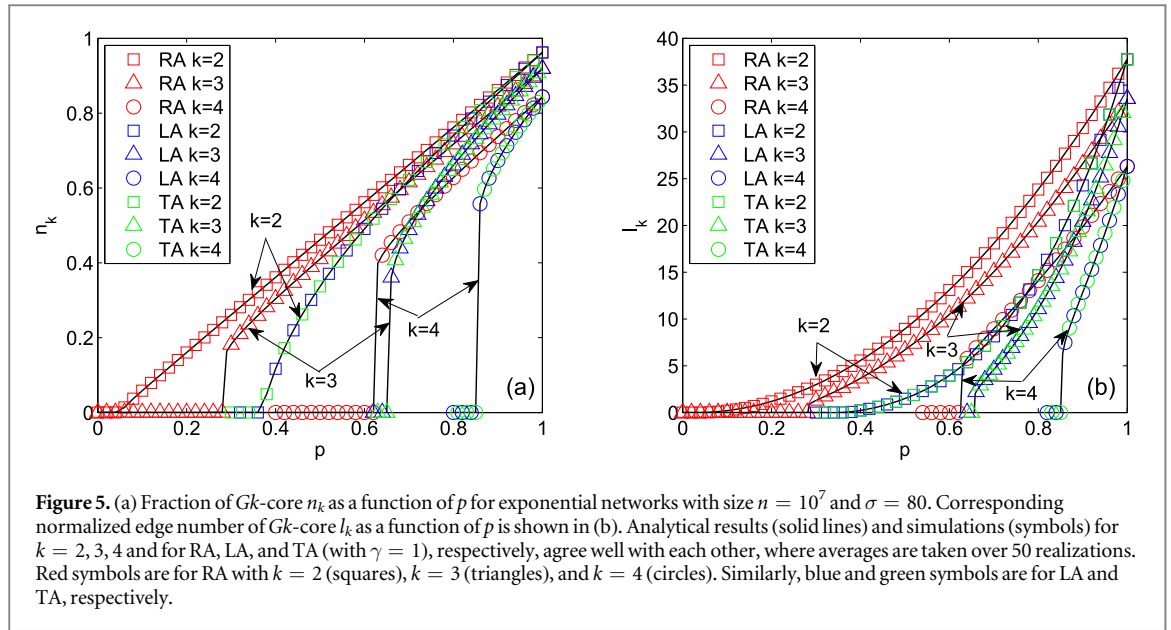
Interestingly, in all cases RA seems to cause more damage to the network than LA does, reminiscent of the giant-component based results observed in [18–20]. The localization effect takes over in LA and leads to a larger  $G_k$ -core for any given  $k$ .

When it comes to  $G_k$ -core stability results under RA and LA shown in figure 4, we find that  $S_k^{\text{LA}} > S_k^{\text{RA}}$  for all  $k$  and  $\ell$ . This means that  $G_k$ -core under LA is more stable than under RA in line with the robustness results. Similar to ER networks, we observe a natural hierarchy that  $S_{k_1} > S_{k_2} > S$  for  $k_1 < k_2$  under all attacks, indicating the less stability of inner  $G_k$ -core with larger  $k$ . It is worth mentioning that  $G_2$ -cores in random regular networks are more stable than in ER networks as  $S_2$  shown in figure 4 is almost level under both RA and LA. This is because random regular networks have a much narrower degree distribution than ER networks even after attacks, which induces less small degree nodes such as 2-leaves stabilizing the  $G_2$ -core.

### 3.3. Exponential networks

An exponential network has the degree distribution  $P(q) = (1 - e^{-1/\sigma})e^{-q/\sigma}$  for  $q \geq 0$  and  $G_0(x) = (1 - e^{-1/\sigma})/(1 - xe^{-1/\sigma})$ . Exponential networks are common in many real-world networks [2, 24, 27] and they have average degree  $e^{-1/\sigma}/(1 - e^{-1/\sigma})$  which is approximately equal to the parameter  $\sigma$  for large  $\sigma$ . Figure 5 shows the behavior of  $n_k(p)$  and  $l_k(p)$  as functions of occupation fraction  $p$  for exponential networks with  $\sigma = 80$  under RA, LA, and TA with  $\gamma = 1$ .





**Figure 5.** (a) Fraction of  $G_k$ -core  $n_k$  as a function of  $p$  for exponential networks with size  $n = 10^7$  and  $\sigma = 80$ . Corresponding normalized edge number of  $G_k$ -core  $l_k$  as a function of  $p$  is shown in (b). Analytical results (solid lines) and simulations (symbols) for  $k = 2, 3, 4$  and for RA, LA, and TA (with  $\gamma = 1$ ), respectively, agree well with each other, where averages are taken over 50 realizations. Red symbols are for RA with  $k = 2$  (squares),  $k = 3$  (triangles), and  $k = 4$  (circles). Similarly, blue and green symbols are for LA and TA, respectively.

We first observe from figure 5, somewhat surprisingly, that the effect of an LA is the same as that of an TA (with  $\gamma = 1$ ) for the exponential network under consideration. Unlike the ER network case, this ‘equivalence’ nevertheless is approximate as we can prove  $n_k^{\text{LA}}(p) = n_k^{\text{TA}}(p)$  and  $l_k^{\text{LA}}(p) = l_k^{\text{TA}}(p)$  if

$$t = \frac{p - 1 + e^{-\frac{1}{\sigma}}}{pe^{-\frac{1}{\sigma}}} = 1, \quad (23)$$

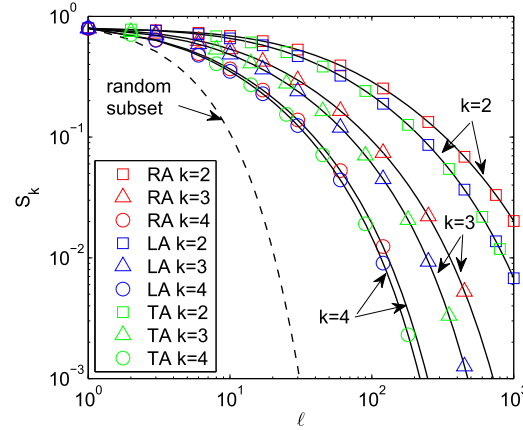
where  $t = G_\gamma^{-1}(p)$  is defined in section 2.3, by using standard argument with (13), (14), (20), (21). The equality of (23) holds precisely when  $\sigma \rightarrow \infty$ , implying a sufficiently dense network. The phenomenon that  $n_k^{\text{RA}}(p) > n_k^{\text{LA}}(p) \approx n_k^{\text{TA}}(p)$  and  $l_k^{\text{RA}}(p) > l_k^{\text{LA}}(p) \approx l_k^{\text{TA}}(p)$  can be understood as there are more hub (i.e. high-degree) nodes in exponential networks than ER networks. The degree heterogeneity effect in LA, on one hand, accelerates the fragmentation process making the exponential network more fragile under LA than RA (as compared to the case of ER networks in figure 1), and on the other hand, elegantly balances the TA in exponential networks in terms of the number and edge of  $G_k$ -cores.

When it comes to the  $G_k$ -core percolation phase transition, similar qualitative phenomenon appears for exponential networks: there are continuous phase transitions for  $k = 2$  and discontinuous phase transitions for  $k \geq 3$  regardless of the attack scenarios. This also highlights the distinction between vaguely defined heavy-tailed distributions and the difficulty in distinguishing between them [28]. For example, clear critical phenomenon is observed here for exponential networks with average degree around 80 here while there is only a trivial  $G_k$ -core for power-law networks irrespective of their connectedness [11].

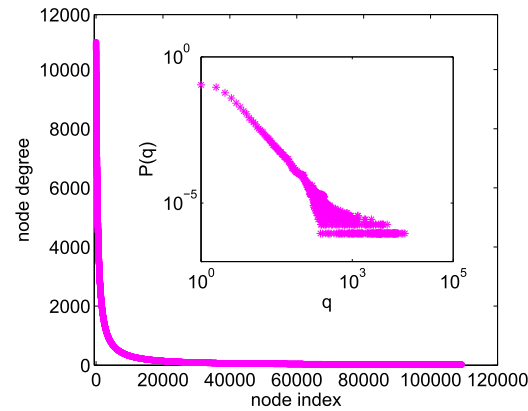
A phenomenon that is worthy of noting is the breakup of the monotonicity of  $n_k(p)$  and  $l_k(p)$  with respect to  $k$  in exponential networks. For example, there is a marked hierarchy for example  $n_2(p) > n_3(p) > n_4(p)$  in ER networks irrespective of the attack types (see figure 1(a)). However, we observe from figure 5(a) for instance that  $n_3^{\text{RA}}(p) > n_2^{\text{TA}}(p)$  under a moderate attack with, say,  $0.3 \leq p \leq 0.6$ . This results from the degree heterogeneity in exponential networks, in which the TA targeting at hubs can produce a much smaller  $G_2$ -core than a  $G_3$ -core under RA, underscoring the cohesive role of hubs in the network robustness.

Next, in figure 6 we display the relationship between  $G_k$ -core stability of exponential networks under RA, LA, and TA when the fraction of  $G_k$ -core is fixed at 0.8, the same as in figure 2 for ER networks and in figure 4 for random regular networks. As one expects, all  $S_k^{\text{RA}}$ ,  $S_k^{\text{LA}}$ , and  $S_k^{\text{TA}}$  decay slower than the random subset case with outer  $G_k$ -cores (namely, smaller  $k$ ) more stable since fewer nodes are removed for smaller  $k$ . Also, as a result of (23) together with (15) and (22),  $S_k^{\text{LA}}$  and  $S_k^{\text{TA}}$  overlap in figure 6, meaning that LA and TA have literally the same effect to  $G_k$ -core stability in our exponential networks.

We mention that the exponential degree distribution decays much more rapidly than a power-law and accommodates much more low degree nodes than an ER network does. In our case for example there are about 1.2% 2-leaves in an exponential network, which are nearly 30 times more than those in an ER network. Therefore, the imagined stabilizing effect of ‘anchor nodes’ tends to be negligible and the many low degree nodes may make the  $G_k$ -core unstable (esp. for small  $k$ ) regardless of the attack scenarios. This generally explains our observation when comparing figure 6 with figure 2.



**Figure 6.** Stability of  $Gk$ -core  $S_k$  with relative size 0.8 as a function of  $\ell$  in the log–log format for exponential networks with size  $n = 10^7$  and  $\sigma = 80$ . Analytical results (solid lines) and simulations (symbols) for  $k = 2, 3, 4$  and for RA, LA, and TA (with  $\gamma = 1$ ), respectively, agree well with each other, where averages are taken over 50 realizations. Red symbols are for RA with  $k = 2$  (squares),  $k = 3$  (triangles), and  $k = 4$  (circles). Similarly, blue and green symbols are for LA and TA, respectively. Dashed line corresponds to stability,  $S$ , of a randomly chosen subset with relative size 0.8.

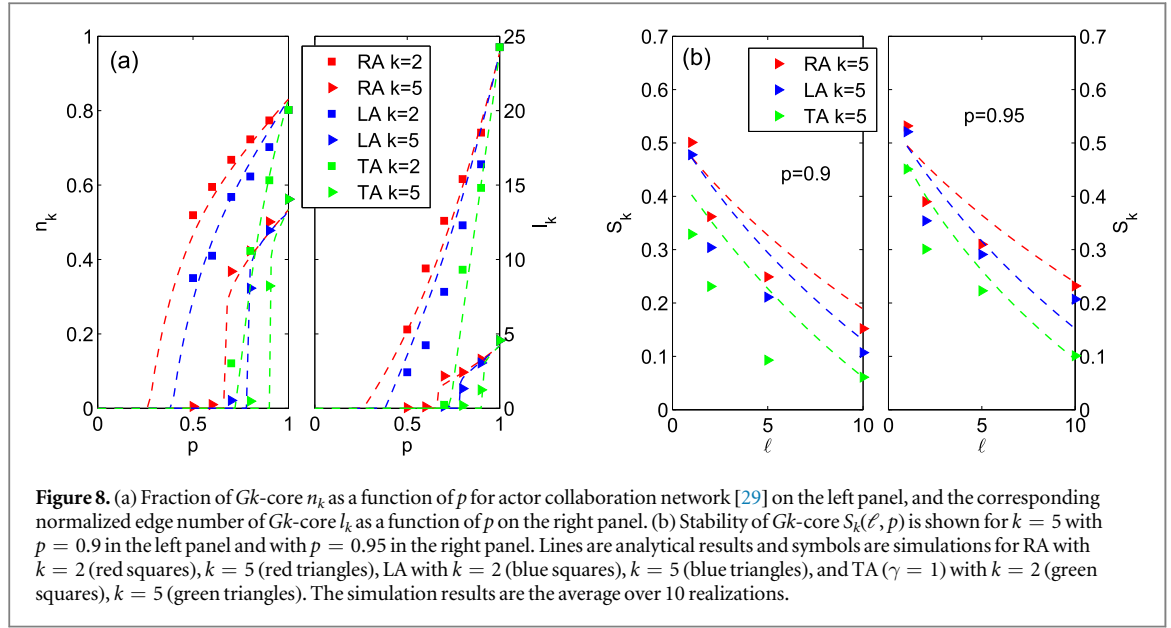


**Figure 7.** Degree distribution of the actor collaboration network, in which nodes are ranked decreasingly with respect to their degrees. Inset: degree distribution in the log–log scale.

#### 4. A real-world network example

We compare the attack robustness and stability of a real-world network against RA, LA and TA by using an actor collaboration network constructed from IMDb in the year 2004 [29, 30]. The network has 1092431 nodes representing movie actors, and 56263702 edges with two actors sharing an edge between them if they ever played in a movie together. The degree distribution is skewed as shown in figure 7. We present the relative size and normalized number of edges in  $Gk$ -cores under RA, LA, and TA in figure 8(a) and stability of  $Gk$ -cores in figure 8(b). More data are collected in tables 2 and 3.

From figure 8(a) and table 2 we observe that that TA is the most harmful attack and RA is the mildest one in all the three types of attacks on the actor collaboration network. For example, when 10% nodes are removed from the network,  $n_k^{TA} < n_k^{LA} < n_k^{RA}$  and  $l_k^{TA} < l_k^{LA} < l_k^{RA}$  for both  $G2$ -core and  $G5$ -core. The stability displayed in figure 8(b) does not have a good approximation between analytical and simulation results, which reveals that the structural features such as degree correlations and clustering may have a non-negligible effect on  $Gk$ -core stability. Nevertheless, figure 8(b) and table 3 allow us to have a tangible understanding on how stable the  $Gk$ -cores are:  $Gk$ -core under TA exhibits the lowest stability, while under RA it exhibits the highest. For example, more than 56% (i.e. 0.436/0.774) nodes that constitute the  $G2$ -core after an RA will remain in the  $G2$ -core after 10 independent repetitions of such an RA. This number decreases to about 42% (i.e. 0.297/0.702) for LA and further to about 24% (i.e. 0.145/0.613) for TA, supporting our theoretical results for networks with heterogeneous degree distributions.



**Table 2.** The fraction  $n_k(p)$  and normalized number of edges  $l_k(p)$  for  $G_k$ -cores in the actor collaboration network under RA, LA, and TA with  $\gamma = 1$ .

	$n_k(p = 1)$		$n_k(p = 0.9)$		$l_k(p = 1)$		$l_k(p = 0.9)$	
	$k = 2$	$k = 5$	$k = 2$	$k = 5$	$k = 2$	$k = 5$	$k = 2$	$k = 5$
RA	0.802	0.562	0.774	0.501	24.270	4.563	18.525	3.299
LA	0.802	0.562	0.702	0.478	24.270	4.563	16.403	3.046
TA( $\gamma = 1$ )	0.802	0.562	0.613	0.329	24.270	4.563	14.814	2.440

**Table 3.** The stability  $S_k(\ell, p)$  for  $G_k$ -cores in the actor collaboration network at  $p = 0.9$  under RA, LA, and TA with  $\gamma = 1$ .

	$S_{k=2}(\ell, p = 0.9)$				$S_{k=5}(\ell, p = 0.9)$			
	$\ell = 1$	$\ell = 2$	$\ell = 5$	$\ell = 10$	$\ell = 1$	$\ell = 2$	$\ell = 5$	$\ell = 10$
RA	0.774	0.690	0.582	0.436	0.501	0.362	0.249	0.152
LA	0.702	0.563	0.409	0.297	0.478	0.304	0.211	0.107
TA( $\gamma = 1$ )	0.613	0.418	0.265	0.145	0.329	0.231	0.093	0.061

## 5. Conclusion

In summary, we have studied the robustness and stability of  $G_k$ -cores of uncorrelated random networks with arbitrary degree distribution. We develop a theoretical framework to systematically gauge network robustness in terms of the relative size and normalized number of edges of  $G_k$ -core under RA, LA, and TA. It is found that continuous phase transition only exists in  $G_2$ -core for all the three types of attacks, and discontinuous transitions are determined for  $G_k$ -core with  $k \geq 3$  in all scenarios. We introduce the  $G_k$ -core stability and show how different types of attacks affect the stability of  $G_k$ -core. Similarities behind the organizing principles underpinning attack robustness and stability are identified, but they are by no means substitutable especially for heterogeneous networks. Methods presented in this work hold promise for more implications in the design and reinforcement of resilient networked systems.

## Acknowledgments

The author would like to thank the reviewers for careful reading and valuable comments. This work was supported by NNSFC (11505127) and Northumbria University.

## ORCID iDs

Yilun Shang  <https://orcid.org/0000-0002-2817-3400>

## References

- [1] Cohen R and Havlin S 2010 *Complex Networks: Structure, Robustness and Function* (Cambridge: Cambridge University Press)
- [2] Barabási A L 2016 *Network Science* (Cambridge: Cambridge University Press)
- [3] Newman M E J 2018 *Networks* 2nd edn (Oxford: Oxford University Press)
- [4] Di Muro M A, Valdez L D, Stanley H E, Buldyrev S V and Braunstein L A 2019 *Phys. Rev. E* **99** 022311
- [5] Dorogovtsev S N, Goltsev A V and Mendes J F F 2006 *Phys. Rev. Lett.* **96** 040601
- [6] Cellai D, Lawlor A, Dawson K A and Gleeson J P 2011 *Phys. Rev. Lett.* **107** 175703
- [7] Morone F, Ferraro G D and Makse H A 2019 *Nat. Phys.* **15** 95–102
- [8] Liu Y Y, Csóka E, Zhou H and Pósfai M 2012 *Phys. Rev. Lett.* **109** 205703
- [9] Azimi-Tafreshi N, Dorogovtsev S N and Mendes J F F 2013 *Phys. Rev. E* **87** 032815
- [10] Jia T and Pósfai M 2014 *Sci. Rep.* **4** 5379
- [11] Azimi-Tafreshi N, Osat S N and Dorogovtsev S N 2019 *Phys. Rev. E* **99** 022312
- [12] Bianconi G 2017 *Phys. Rev. E* **96** 012302
- [13] Kitsak M, Ganin A A, Eisenberg D A, Krapivsky P L, Krioukov D, Alderson D L and Linkov I 2018 *Phys. Rev. E* **97** 012309
- [14] Albert R, Jeong H and Barabási A L 2000 *Nature* **406** 378
- [15] Callaway D S, Newman M E J, Strogatz S H and Watts D J 2000 *Phys. Rev. Lett.* **85** 5468–71
- [16] Newman M E J 2002 *Phys. Rev. E* **66** 016128
- [17] Shang Y 2014 *Phys. Rev. E* **90** 032820
- [18] Shao S, Huang X and Stanley H E 2015 *New J. Phys.* **17** 023049
- [19] Shang Y 2016 *Sci. Rep.* **6** 30521
- [20] Dong G, Xiao H, Wang F, Du R, Shao S, Tian L, Stanley H E and Havlin S 2019 *New J. Phys.* **21** 013014
- [21] Huang X, Gao J, Buldyrev S V, Havlin S and Stanley H E 2011 *Phys. Rev. E* **83** 065101(R)
- [22] Shang Y 2018 *J. Stat. Phys.* **170** 141–64
- [23] Baxter G J, Timár G and Mendes J F F 2018 *Phys. Rev. E* **98** 032307
- [24] Newman M E J, Strogatz S H and Watts D J 2001 *Phys. Rev. E* **64** 026118
- [25] Shang Y 2019 *IEEE Trans. Syst. Man Cybern. Syst.* **49** 821–32
- [26] Yuan X, Dai Y, Stanley H E and Havlin S 2016 *Phys. Rev. E* **96** 062302
- [27] Cimini G, Squartini T, Saracco F, Garlaschelli D, Gabrielli A and Caldarelli G 2019 *Nat. Rev. Phys.* **1** 58–71
- [28] Broido A D and Clauset A 2019 *Nat. Commun.* **10** 1017
- [29] Boldi P and Vigna S 2004 *Proc. 13th Int. World Wide Web Conf.* (New York: ACM Press) pp 595–601
- [30] Rossi R A and Ahmed N K 2015 *Proc. 29th AAAI Conf. on Artificial Intelligence* (Palo Alto, CA: AAAI Press) pp 4292–93